



Digitized by the Internet Archive
in 2019 with funding from
University of Alberta Libraries

<https://archive.org/details/Ross1965>

Thesis
1965
170

THE UNIVERSITY OF ALBERTA

A COMPARATIVE STUDY OF INTRAClass GROUPING AND NON-GROUPING
IN GRADE SEVEN ACHIEVEMENT

by

DOUGLAS HEBER ROSS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF EDUCATION

DEPARTMENT OF EDUCATIONAL ADMINISTRATION

EDMONTON, ALBERTA

APRIL, 1965

UNIVERSITY OF ALBERTA
FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "A Comparative Study of Intraclass Grouping and Non-Grouping in Grade Seven Achievement" submitted by Douglas Heber Ross in partial fulfilment of the requirements for the degree of Master of Education.

ABSTRACT

The problem was to discover whether grade seven pupils achieve significantly higher reading results when they are grouped and given differentiated reading instruction, than pupils who are not grouped.

Growths in reading achievement of two independent samples of grade seven pupils were compared. The testing period covered seven months. In one sample, the pupils were placed in one of three groups within their classrooms for reading instruction. They were grouped according to their reading ability as indicated by the Iowa Silent Reading Test. An attempt was made by the teachers of the grouped classes to differentiate the instruction and materials so that each group would receive instruction suited to their ability level. In the other sample the pupils were not grouped and they received class-as-a-whole instruction.

The two samples were comparable in mental ability and reading ability as measured by the Otis Quick Scoring Mental Ability Test and the Iowa Silent Reading Test at the beginning of the testing period. From the evidence available, it was assumed that the two samples were also comparable in socio-economic background and reading background. The variable of teacher competence was assumed to be controlled because of the large number of teachers involved.

Comparisons were made between the growths as shown by the Iowa Silent Reading Test, the sub-test "comprehension" and the sub-test "word-meaning."

The statistical tests indicated that:

1. There was no significant difference in the overall growth of achievement between the grouped and the ungrouped students.
2. There was a tendency for the students in the above-average grouped classes to gain relatively less than their counterparts.
3. There was no significant difference in the overall growth of achievement between the below-average samples.
4. There was a significant difference in the growth of the reading skill "comprehension" in favor of the below-average grouped students as compared to their counterparts.
5. There was no significant difference in the growth of the reading skill "word meaning."
6. In both the grouped and ungrouped classes the range in reading ability tended to diminish.

The evidence seemed to indicate that on the whole, the type of grouping used for this study produced no difference in achievement as compared to non-grouping. The one exception favored the below-average grouped students. Other studies on grouping have also found that the benefits, if any, usually accrue to the low-ability, grouped students.

ACKNOWLEDGEMENTS

The investigator acknowledges, with appreciation, the help given by Mr. R. I. McLoughlin, Curriculum Consultant, and Mr. H. Toews and Mr. H. C. Rhodes, Superintendents of Schools. Gratitude, for their co-operation, is expressed to the principals of the schools involved in the study.

Appreciation is also expressed to Dr. W. D. Knill, Dr. F. Enns, and Dr. S. Sigurdson for constructive counsel and assistance.

TABLE OF CONTENTS

CHAPTER	PAGE
I. THE PROBLEM AND DEFINITIONS OF TERMS USED	1
The Problem	1
Statement of the problem	2
Significance of the study	3
Definitions of Terms Used	3
Assumptions	4
Delimitations	9
Limitations	10
Procedure	10
Null Hypotheses	12
II. REVIEW OF THE LITERATURE	13
Early Experiments	13
Later Research on Intraclass Grouping	20
Later Research on Interclass Grouping	25
Individualized vs. Group Reading	30
Acceleration and Enrichment	34
Summary	37
III. THE DESIGN OF THE STUDY	39
Selection of Sample and Testing Procedure	42
A Comparison of Grouped and Ungrouped Students	47
Background	47
Intelligence	47

CHAPTER	PAGE
The Testing Instruments	48
Statistical Treatment	52
IV. ANALYSIS OF THE DATA	54
Comparison of Overall Growth on Total Score	54
Comparison of Above-Average Students on Total Score	55
Comparison of Below-Average Students on Total Score	58
Comparison of Gains on Sub-Test, "Comprehension"	61
Comparison of Gains on Sub-Test, "Word-Meaning"	65
Comparison of Variances	69
Summary of Results	69
V. CONCLUSIONS AND IMPLICATIONS	71
Hypothesis I	71
Hypothesis II	72
Hypothesis III	73
Hypothesis IV	73
Hypothesis V	74
Hypothesis VI	75
Summary of Conclusions	76
Implications of the Study	77
Recommendations for Further Study	78
BIBLIOGRAPHY	79

LIST OF TABLES

TABLE		PAGE
I.	Frequency Distributions of Reading Scores on Pre-Test Before Removal of Twenty-nine Students	44
II.	Frequency Distributions of I.Q.'s of Each Sample	49
III.	I.Q. Distributions of Above-Average and Below-Average Students	50
IV.	Total Reading Scores on Pre-Test and Post-Test by Students in Both Samples	56
V.	Total Reading Scores on Pre-Test and Post-Test by Above-Average Pupils	59
VI.	Total Reading Scores on Pre-Test and Post-Test by Below-Average Pupils	60
VII.	Comprehension Scores on Pre-Test and Post-Test by All Students in Both Samples	62
VIII.	Comprehension Scores on Pre-Test and Post-Test by Above- Average Pupils	63
IX.	Comprehension Scores on Pre-Test and Post-Test by Below-Average Pupils	64
X.	Total Word Meaning Scores on Pre-Test and Post-Test by Students in Both Samples	66
XI.	Word Meaning Scores on Pre-Test and Post-Test by Above-Average Pupils	67
XII.	Word Meaning Scores on Pre-Test and Post-Test by Below-Average Pupils	68

TABLE	PAGE
XIII. Summary of Tests of Significance	70

LIST OF FIGURES

FIGURE	PAGE
1. Scores of Grouped and Ungrouped Students After Removal of Students Who Missed Post-Test	45
2. Comparison of Overall Student Scores on Pre-Test and Post-Test	57

CHAPTER I

THE PROBLEM AND DEFINITIONS OF TERMS USED

Grouping in the classroom in order to give differentiated instruction is an accepted practice in many Canadian elementary schools, especially for reading instruction. It is also a commonly accepted practice in the same school systems to stop intraclass grouping as soon as these children reach grade seven, and to give all those in the same class identical lessons, no matter what their range in ability. Yet there is just as great a range, if not greater, in the secondary grades as in the elementary, and so perhaps grouping for reading instruction should be continued at the secondary level.

I. THE PROBLEM

In the introduction to the section on developmental reading in the Curriculum Guide for Junior High School Literature for the Province of Alberta, it is stated:

. . .Between twenty and thirty percent of pupils entering the junior high school read so poorly as to carry on their studies with great difficulty. In fact, it is not uncommon to find a reading range of six or more grades among pupils in a single class. This fact suggests to us the importance of considering the individual differences among pupils with a view to providing some differentiated guidance of a corrective or developmental type.¹

The curriculum guide then goes on to suggest how the teacher might

¹Department of Education, Province of Alberta, Curriculum Guide for Junior High School Literature (Edmonton: September, 1951), p. 36.

diagnose the reading difficulties of the individual students in the class and tells how the individual student might be encouraged to read library books suitable to his interest and reading level.² Finally, it mentions the two Reading for Meaning Workbooks, one for Grade VII and one for Grade VIII and admonishes the teacher to make adequate use of these aids. This introduction to the reading section however, as far as the curriculum guide is concerned, ends the Department's responsibility for differentiating instruction in grade seven reading.

The British Columbia Provincial Department of Education subscribes to the idea that grouping within the classroom and giving differentiated instruction is necessary to provide in part for the individual differences found within a heterogeneous class. The programme of studies for that province specifically requires that the teachers do group within the classroom.³

Statement of the problem. Do grade seven pupils achieve significantly higher reading results when they are grouped for reading, and when the teacher gives differentiated instruction? In solving this problem, samples were drawn from the school populations surrounding Dawson Creek, B. C., and Grande Prairie, Alberta.

In attempting to solve the problem, answers to the following questions were sought:

²Ibid., pp. 37-41.

³Division of Curriculum, Department of Education, Province of British Columbia, Programme of Studies for Elementary Schools, Supplement, Grade VII (Victoria: Queen's Printer, 1961), pp. 8-11; 27-34.

Is there a significant difference in the overall growth of reading achievement between the grouped and ungrouped students when the students are compared?

Is there a significant difference in the growth of achievement between the above-average reading groups?

Is there a significant difference in the growth of achievement between the below-average reading groups?

Does grouping increase the range in reading achievement?

Does grouping affect some aspects of reading achievement more than others?

Significance of the study. If it could be shown that grade seven pupils do achieve significantly higher reading results when they are grouped, and given differentiated instruction within these groups, then this would be significant for educational administrators at all levels.

II. DEFINITIONS OF TERMS USED

Grouping. Grouping means the placing of students in one of three reading groups in a class according to reading ability as shown by the Iowa Silent Reading Tests.

Grouped students. This refers to those students in classes where grouping took place for purposes of differentiated instruction in reading. The students of the Dawson Creek area were grouped in this manner.

Ungrouped students. Those students who were not grouped

within their classes for reading instruction and did not receive differentiated instruction. These were the students of the Grande Prairie schools.

Above-average pupils. Those pupils who scored one standard deviation above the mean on the pre-test are designated "above-average." For the grouped students, this included those who scored 170 and over on the Iowa Silent Reading Test, Form Am. For the ungrouped, it included those who scored 168 and over.

Below-average pupils. The below-average pupils are those who scored one standard deviation below the mean on the pre-test. For both the grouped and the ungrouped pupils, it included those who scored 145 and below on the Iowa Silent Reading Test, Form Am.

Differentiated instruction. This means that partially different reading material was used for instructional purposes for the different reading groups as set out in the Programme of Studies for Elementary Schools for the Province of British Columbia.⁴

III. ASSUMPTIONS

The objectives of the reading program for grade seven, for both school systems referred to in this study, were assumed to be comparable. The British Columbia programme of studies for elementary schools interprets its objectives as follows:

⁴Ibid., pp. 8-11; 27-34.

. . .The English programme requires each year not only a review of the work of earlier grades but also a progressively more mature treatment of each item in the programme,. . .Progress from grade to grade thus implies an ever widening and deepening spiral of knowledge and understanding and an increasing mastery of the basic skills.⁵

The curriculum guide for literature for Alberta appears to interpret its objectives similarly.

. . .Today it is recognized that reading is far too complex a psychological process to be mastered by the time a child has completed Grade VI. Instead, it is now agreed that the elementary school merely begins a process which must be continued and expanded in the junior high school, and the college or university. . .The reading program of the junior high school should provide whatever skills, habits, and attitudes that are needed for further growth in reading.⁶

Because neither Figure 1, nor Figure 2, pp. 45 and 57 showed extreme skewness, and because the samples were large, it was assumed that the distributions of both samples approximated normal curves. Walker and Lev state:

An interesting and important fact is that even if the population is not normally distributed, the sampling distribution of the mean may be well approximated by the normal distribution if samples are large (say $N > 30$). Under such circumstances, the distribution of the mean can be treated as normal with mean μ and standard error $\frac{\sigma}{\sqrt{N}}$.⁷

The teachers involved in this study were all given identical written instruction for administering the tests. It was assumed that all the teachers followed these directions and that the testing was the

⁵Ibid., p. 7.

⁶Department of Education, Province of Alberta, op. cit., pp. 36-7.

⁷Helen M. Walker, and Joseph Lev, Statistical Inference (New York: Holt, Rinehart and Winston, Inc., 1953), p. 132.

same for all the students. The answer sheets were all marked and the results recorded by the writer.

The variable of teaching competence was assumed to be controlled because of the large number of teachers involved. There were thirteen teachers of grouped classes and fifteen teachers of ungrouped classes.

The school populations from which the samples were drawn were assumed to have a comparable socio-economic background because of their proximity to each other. Furthermore, both school populations lived mainly in farming communities.

The median salary range for male wage earners in 1951 in both Grande Prairie and Dawson Creek was \$2,000 to \$2,499. Grande Prairie had twenty-five male wage earners earning \$4,000 or over in 1951 while Dawson Creek had twenty-six in this category.⁸

The population of Grande Prairie in 1961 was 8,352. At the same time Dawson Creek's population was 10,946.⁹ Both were incorporated as cities in 1958.

Grande Prairie had thirty-seven elementary classrooms, one large junior high school, one senior high school and one separate school in 1961. Dawson Creek had six elementary schools, two junior high schools, one senior high school and one parochial school in 1963.

Grande Prairie has a 120-bed hospital; Dawson Creek a 100-bed

⁸Dominion Bureau of Statistics, Ninth Census of Canada: 1951: Earnings and Employment of Wage Earners, Vol. V (Ottawa: Queen's Printer, 1953), Table 16-19 and 16-21.

⁹Dominion Bureau of Statistics, Canada Year Book (Ottawa: Queen's Printer, 1963-64), p. 169.

hospital.

The trading area population of both cities was estimated at 35,000 in 1961.

The reading content of the textbooks used by the pupils of both samples was assumed to be comparable.

The length of time spent by both samples in reading was comparable.

It was assumed that the pupils in the two samples had teachers who followed their respective curriculum guides and used the authorized textbooks and materials in instructing their students in the elementary grades.

Based on the above assumption and an inspection of the elementary school curriculum guides for both provinces, a further assumption was made that the students of both samples had a similar background in reading.

For example, in grades one to three, in British Columbia, the Curriculum Foundations series of readers is authorized, and the corresponding Think and Do workbooks may be used but are not prescribed. In Alberta, this same series of readers and workbooks are authorized, although in Alberta two other series of readers are also authorized.

In grades four to six, the Canadian Parade series of readers and the Canadian Reading Development series are authorized for British Columbia. In Alberta, three series are authorized for grades four to six, two of which are identical with those authorized for British Columbia, namely, the Canadian Reading Development series and the Canadian Parade series.

In both provinces, intraclass grouping for reading instruction in grades one to six is encouraged. In both provinces the grade seven pupils take literature and reading. In both cases the reading program is divided into developmental reading and remedial reading. Both programs of study suggest that the Iowa Silent Reading Tests and the Iowa Every-pupil Tests of Basic Skills be given at the beginning of grade seven to determine silent reading ability at this grade level.

Both courses of study encourage extensive and varied free reading and give suggested lists of suitable books for grade seven. A comparison of the two lists reveals a total of thirty-two books which are identical for grade seven. In addition there are five books listed for B.C. grade seven which are listed for Alberta grade eight or nine, and there are nine books listed for B.C. grade seven which are listed for Alberta grade five or six. Thus a total of forty-six library books are listed in the courses of study which could be read by the grade seven students in both provinces. There are approximately 130 listed for B.C. and 220 listed for Alberta, for grade seven. Furthermore, the students are not restricted to the books listed in the courses of study and therefore it is possible that many of the books they read are the same.

Both courses of study stress the importance of continuing oral reading in grade seven and state that the pupils should read orally from their literature books.

Both encourage the use of audio-visual aids as a means of motivation and extension of pupil experience.

The main difference between the courses of study lies then in the

emphasis on grouping and differentiation of instruction in the B.C. course of studies for Grade Seven, as compared to the treatment of the class as a whole in Alberta. In the Alberta course of studies these instructions to the teacher appear:

Say to the class: "We are going to read another poem. The lines are short enough that you will not need to move your eyes many times per line. The story which the poem tells us is an interesting one about a hero who helped to save his country's fleet from destruction. Turn to page 204 and begin reading."¹⁰

After reading, the teacher asks six questions of the whole class.

By contrast, the B.C. course of studies states that students should be assigned to three reading groups within each class. It also states that those pupils scoring a 6.5 grade level or higher should be given developmental instruction and those scoring less should be given remedial instruction. A system of letters beside the reading selections, A for above-average, B for average and C for below-average, indicate to the teacher which selections should be given to each of the three reading groups during the year.

IV. DELIMITATIONS

The study was delimited to the pupils that attend school in the cities of Dawson Creek and Grande Prairie and their respective areas. Schools that had fewer than twenty-four grade seven pupils were not included. It was further delimited to grade seven reading. The time covered a period of approximately seven months.

¹⁰Department of Education, Province of Alberta, Curriculum Guide for Junior High School Literature, op. cit., p. 43.

V. LIMITATIONS

The Otis Quick Scoring Mental Ability Test is limited to its ability to measure verbal competence, but it does this well and it does indicate scholastic aptitude.¹¹

The statistical comparisons of the sub-tests on the Iowa Silent Reading Test were limited to "comprehension" and "word meaning." To have compared achievement on each of the other six sub-tests in terms of total scores, above-average scores and below-average scores would have added thirty-six statistical comparisons to the twenty-one that were made. This would have added unnecessary bulk to the study. The purpose of the study was to determine whether grouping produced greater achievement than non-grouping. Therefore, although there were eight sub-tests altogether, the two, "comprehension" and "word meaning," were selected as being the most indicative of reading ability. They were also intended to give an indication of whether the results of the sub-tests were essentially in agreement with the results of the overall test.

VI. PROCEDURE

Nine-hundred of the Otis Quick Scoring Mental Ability Tests Gamma, Form Em, and of the Iowa Silent Reading Tests, Form Am, were mailed to the superintendents of the two school systems involved three weeks in advance of the pre-test. The superintendents distributed the tests, the answer sheets, the manuals and the investigator's instructions, to each

¹¹Oscar Krisen Buros (ed.), The Fifth Mental Measurements Year-book (Highland Park, New Jersey: The Gryphon Press, 1959), p. 499.

of their teachers. The tests were administered by the teachers on October 22, 1963. The mental ability tests were given on the morning of October 22, and the reading tests on the afternoon of the same day. The test booklets were collected by the principals and kept on file. The answer sheets were collected by the principals immediately after the testing and mailed to the investigator for scoring. Seven months later, on April 20, 1964, the same reading tests were administered as a post-test. The answer sheets were again collected by the principals and mailed directly to the investigator for scoring.

The mental ability test was given at the same time as the reading pre-test so that students who had I.Q.'s which were extremely different from the other students' I.Q.'s could be removed if it were necessary.

The silent reading test results were used to assign pupils to one of the three reading groups within the grouped classes. It was also used as a pre-test and as a post-test.

The data from these tests were used as follows:

1. The means and variances of both samples were compared on the total test.
2. The means and variances of the above-average scores in each sample were compared on the total test.
3. The means and variances of the below-average scores in each sample were compared on the total test.
4. The means and variances of both samples were compared on the sub-test "comprehension."
5. The means and variances of the above-average scores in each sample were compared on the sub-test "comprehension."

6. The means and variances of the below-average scores in each sample were compared on the sub-test "comprehension."

7. The means and variances of both samples were compared on the sub-test "word meaning."

8. The means and variances of the above-average scores in each sample were compared on the sub-test "word meaning."

9. The means and variances of the below-average scores in each sample were compared on the sub-test "word meaning."

VII. NULL HYPOTHESES

The following null hypotheses were formulated before applying tests of significance in order to answer the questions posed in Chapter I, p. 3:

1. There is no significant difference in the overall growth of reading achievement between the grouped and the ungrouped students.

2. There is no significant difference in the growth of achievement between the above-average reading groups.

3. There is no significant difference in the growth of achievement between the below-average reading groups.

4. There is no significant difference in the growth of the reading skill, "comprehension."

5. There is no significant difference in the growth of the reading skill, "word meaning."

6. There is no significant difference between any of the variances which are compared.

CHAPTER II

A REVIEW OF THE LITERATURE

Grouping the pupils for instruction within a school can take many forms. There appear to be two basic methods of grouping--intraclass and interclass. With intraclass grouping, the pupils of a given class remain together in the same classroom, but are grouped homogeneously in two or more groups. With interclass grouping, the pupils within a school are placed in homogeneous groups, and each group is placed in a different classroom.

Variations of these two basic arrangements are due, in part, to three basic methods of selecting the groups. Selection can be done on the basis of ability, achievement, or interest, or combinations of these three.

Still further variations occur depending on whether or not instruction and materials are changed to suit the various groups.

Since the type of organization used in this study was intraclass grouping, the review will deal with intraclass studies first. A discussion of the literature on interclass grouping, a brief review of individualized versus group reading, and finally, a brief review of acceleration and enrichment will follow.

I. EARLY EXPERIMENTS

This investigation attempted to discover whether an A, B, C type of intraclass grouping was superior to non-grouping. The fact that

educators have been studying this type of organization, among others, in their concern with solving the problem of individual differences since the beginning of the century, is shown by a review of the literature on this problem.

Passow, after reviewing early research, concluded that:

Although the practice of grouping students reached its peak in the 1920's and 1930's, the origins of grouping go back into the last century. . . .The Santa Barbara Concentric Plan was begun after the turn of the century, with each grade divided into A, B, and C sections, each mastering the same fundamentals for each subject, but the A's doing more extensive work than the B's, and the B's more than the C's. . . .However, it was not until 1916 that any serious attempt was made to study homogeneous grouping with something resembling controlled experimentation.¹

Thus it can be seen that although this method of organization was tried out fifty years ago, there is no proper indication of its success or failure because it was not properly evaluated at the time.

Although many of the studies done on grouping in the 1920's were simply reports on what was then a relatively new method of organization, a few made some attempt to assess their grouping practices on an experimental basis. One such study, by Worlton, had results which conflict with many of the later studies.² Worlton's "experimental" study was conducted during the second semester of the school year 1923-24. Worlton said that he attempted to control as many of the variables which influence pupil progress as possible. The pupils were grouped homogeneously on the basis of their ranks "on tests used for classification

¹Harry A. Passow, "The Maze of Research on Ability Grouping," Educational Forum, 26:282, March, 1962.

²J. T. Worlton, "The Why of Homogeneous Classification," Elementary School Journal, 27:265-74, 1926.

purposes." He did not say whether instruction was differentiated, nor did he state whether the homogeneous classes were organized on an intra-class or interclass basis.

He found that the pupils in the homogeneous classes made greater gains in knowledge of subject matter than the pupils in the heterogeneous classes.³ Worlton appeared to have been convinced that homogeneous grouping was beneficial for students.

Another early study, by Barthelmess and Boyer, also found that homogeneous grouping produced significantly higher results in achievement for the grouped students.⁴ They do not state whether intraclass or interclass grouping was used. Their study covered the period from September 1930 to January 1932, in five schools in Philadelphia. The experimental conditions appear to have been reasonably well controlled. The pupils were grouped by I.Q. and in most cases each grade had three groups. Minimum courses for the "bright" and "dull" groups in arithmetic, history and geography were given, and suggestions were made to the teachers of the grouped classes on how to adjust their methods to fit the high and low groups. It is not stated what form this adjustment took. Control groups of heterogeneous classes were set up, and matched with the homogeneous groups on the basis of grade placement, I.Q., C.A., and "efficiency" of teachers. The comparison was made between two sets of grade four pupils.

³Ibid., p. 267.

⁴Harriet Barthelmess and P. Boyer, "An Evaluation of Ability Grouping," Journal of Educational Research, 26:284-94, December, 1932.

Significant differences were found in favor of the grouped classes in three of the five schools. No significant difference was found in one school while the other school reported a significant difference in favor of the heterogeneous classes. On the total scores for all five schools there was a significant difference in favor of the grouped classes. The authors concluded that their study offered exceedingly strong evidence that homogeneous grouping can be a factor in securing improvement in certain skill subjects.

The findings of the present study are quite similar to an experiment reported by Billet in 1928. Miller and Otto refer to Billet's study.⁵ It included 109 pupils in ninth grade English. The data showed that in four cases out of five the difference of the mean gains was in favor of the slow homogeneous group. In four cases out of five the difference of the mean gains was in favor of the heterogeneous group. Miller and Otto felt that the differences probably were not significant, but they added that a repetition of the experiment the following year corroborated the original findings.

Many of the studies on grouping have found no significant differences in academic achievement between homogeneous and heterogeneous groups. Because of conflicting findings with other studies, however, it has still to be proved that homogeneous grouping under proper conditions, would be more beneficial than heterogeneous grouping. Such a sentiment was expressed by one of the early researchers. Passow referred to a

⁵W. S. Miller and Henry J. Otto, "Analysis of Experimental Studies in Homogeneous Grouping," Journal of Educational Research, 21:100, May, 1930.

critical review of the research on ability grouping made by Rock, in 1929.⁶ He said that Rock concluded:

The experimental studies of grouping which have been considered, fail to show consistent, statistically or educationally significant differences between the achievement of pupils in homogeneous groups and pupils of equal ability in heterogeneous groups. This failure to realize one of the important advantages claimed for ability grouping is not, however, evidence that homogeneous grouping cannot result in increased academic achievement. Neither do the experiments show that other claims made for grouping cannot be attained under proper organization.

By 1936 however some educators were wondering whether grouping was, in itself, enough. Cornell, in a chapter of the Thirty-Fifth Yearbook, came to this conclusion:

The results of ability grouping seem to depend less upon the fact of grouping itself than upon the philosophy behind the grouping, the accuracy with which grouping is made for the purposes intended, the differentiations in content, method and speed, and the technique of the teacher, as well as upon more general environmental influences. Experimental studies have in general been too piecemeal to afford a true evaluation of the results, but when attitudes, methods, and curricula are well-adapted to further adjustment of the school to the child, results, both objective and subjective, seem favorable to grouping.⁷

Wilhelms and Westby, in an article called "Grouping Research Offers Leads," arrived at these "undeniable findings" about the old-style A, B, C grouping:

1. It does not produce anywhere near the hoped for homogeneity (except in the one subject area, or I.Q., where the specific grouping is done) because individuals are not homogeneous within themselves.
2. . . .It does not produce more than microscopic improvements,

⁶Passow, op. cit., p. 283.

⁷Ethel L. Cornell, "Effects of Ability Grouping Determinable from Published Studies," The Grouping of Pupils, Thirty-Fifth Yearbook of the National Society for the Study of Education (Bloomington, Illinois: Public School Publishing Company, 1936), p. 304.

if these, in subject matter mastery.⁸

They then suggest other ways of meeting the individual needs of students such as non-graded elementary schools, individualized reading programs, team teaching, team teaching coupled with large and small group instruction, and guidance to help reveal special abilities or needs.

As early as 1930, Miller and Otto, in criticizing twenty of the earlier studies on grouping, dating from 1920 to 1929, found the evidence from most of the studies to be unreliable because of poor experimental design.⁹ Most of the studies lacked a control group. Many were evaluated by such methods as noting the change in the number of failures, or by comparing teachers' marks. From a review of the nine studies which used control groups and objective tests however, Miller and Otto were able to draw some conclusions. They state:

The tendency all too often is to use exactly the same method for the different sections. Unfortunately, we do not yet know enough about differences between methods of instruction for the different levels of ability to make proper adaptations. If it is true that in the conventional school to-day the achievement of the slow pupils corresponds much more closely with their mental level than does that of the bright pupils, then our present methods must be quite efficacious for the dull pupil. Possibly the future may bring equally efficient instruction for the bright pupil.

Although these words were written over thirty years ago, they seem to apply almost equally well to the present day. Possibly we still do not know enough about differences between methods of instruction for

⁸F. T. Wilhelms and Dorothy Westby-Gibson, "Grouping Research Offers Leads," Educational Leadership, 18-412, April, 1961.

⁹Miller and Otto, op. cit., p. 101.

the different levels of ability to make proper adaptations. It appears that this may have been the case for the two school systems observed in the present study. It is also noteworthy that the same difficulty seems to exist in one of the larger Alberta cities where all the elementary pupils in the public school system are grouped in a non-graded type of organization. A recent study of this organization was undertaken by Melnychuk.¹⁰ He concluded that although the pupils are grouped homogeneously for instruction, "there is considerable evidence to indicate that many teachers use the same techniques and material for each group that they teach."

Elsewhere in Alberta, it has been found that grouping rarely takes place in the classroom. Even where it does, there is seldom a change in the curriculum. Worth states:

Elementary school workers have long paid lip-service to making provisions for individual differences among pupils. But provisions for accommodating these differences have traditionally consisted of little more than a grouping together for instructional purposes of children of like age, ability, and/or achievement with few real changes in curriculum. Even this minimum provision is relatively rare in Alberta schools, according to a recent survey of grouping practices.¹¹

Worth is referring to Arbeau's study.¹²

After more than thirty years it is still to be hoped that the

¹⁰Rudolph Steve Melnychuk, "Academic Achievement of Pupils in the Edmonton Continuous Progress Plan" (unpublished Master's thesis, The University of Alberta, Edmonton, 1964), p. 89.

¹¹Walter H. Worth, "Elementary Education--The Critical Years," Alberta School Trustee, 34:26, December, 1964.

¹²A. M. Arbeau, "A Survey of Pupil Grouping Practices in Grade I to XII in Alberta Schools" (unpublished Master's thesis, The University Of Alberta, Edmonton, 1963).

future may bring equally efficient instruction for all pupils.

II. LATER RESEARCH ON INTRACLASS GROUPING

Many writers have indicated that for grouping to be of any value, the different groups should have instruction fitted to their needs. An experiment was conducted in 1940 by Guiler and Edwards to test such a hypothesis.¹³ The experiment was conducted with pupils of grade seven and eight in computational arithmetic. The experimental pupils received instruction based on a diagnosis of their individual difficulties. The control groups pursued the regular course work. Each teacher of an experimental class was given a chart which indicated the weaknesses needing attention for each individual pupil in his class. The teachers of the control groups gave their pupils identical lessons and neither pupils nor teachers of these classes were made aware of individual pupils' difficulties. In the experimental classes, lessons were given to the whole class in areas where the majority of the class showed weaknesses, but when a limited number of pupils had difficulty in a certain area special lessons and practice were given to the pupils concerned. Each pupil was given a workbook with the materials arranged so that the pupils could obtain self-teaching and extensive practice in skills in which they were weak. In all cases the experimental classes achieved significantly higher than the control classes.

Guiler and Edwards concluded from the evidence that group

¹³W. S. Guiler and Vernon Edwards, "An Experimental Study of Methods of Instruction in Computational Arithmetic," Elementary School Journal, 43:353-60, February, 1943.

instruction based on individual diagnosis is more effective than conventional instruction in the improvement of computational habits.

Russell reported an evaluation of a combined interclass, intraclass type of organization used in grades four, five and six.¹⁴ The pupils were grouped homogeneously and changed rooms for instruction in reading. In most classes there were two sub-groups. The achievement of these 278 children was compared with that of 248 children who were not grouped. The two sets of children were matched by C.A., M.A., and reading achievement. At the end of two years, standardized tests revealed no significant differences in reading achievement between the two sets.

Many studies have found that the greatest benefits, if any, accrue to the lower achievers. An experiment was reported by Jones, in which the growth achieved by a control group in a traditional class was compared with the growth of an experimental class. Intraclass grouping was used and instruction was given to meet individual needs.¹⁵ The control group pursued the traditional course of study in reading, arithmetic and spelling, with only minor and incidental treatment of individual problems. The experimental group pursued a curriculum designed to meet individual needs as to level and rate of progress. Two-hundred and fifty grade four pupils were involved.

In order to make statistical comparisons of the growth, Jones

¹⁴D. H. Russell, "Inter-Class Grouping for Reading Instruction in the Intermediate Grade," Journal of Educational Research, 39:62-70, February, 1946.

¹⁵Daisy M. Jones, "An Experiment in Adaptation to Individual Differences," Journal of Educational Psychology, 39:257-71, May, 1948.

placed the pupils in three groups on the basis of I.Q.--those with 110 and above, between 90 and 109, and those below 90. Since the pupils were grouped for instruction mainly on the basis of achievement, Jones really found what the effect of grouping by reading achievement had on the pupils in the three intelligence classifications.

Overall it was found that the experimental group gained two and a half months more than the control group. This difference was significant at the .01 level. Jones also found that in no case was the difference significant for the superior pupils but all differences for the average and below-average pupils were significant. The lower the intelligence classification, the greater was the significance of the difference.

Her conclusions support those of many other writers.

An experimental study was done in 1956 by Holmes and Harvey in which two kinds of intraclass grouping, "flexible," and "permanent," were compared.¹⁶ In permanent grouping, the teacher divided his class into two or more groups at the beginning of the year and little or no shifting was done during the year. In flexible grouping, the teacher taught the class as a whole in introducing new material and then divided the class into groups on the basis of accomplishment in that particular topic.

They concluded that the method of grouping used in organizing these arithmetic classes made very little difference with regard to pupil achievement.

¹⁶Darrell Holmes and Lois Harvey, "An Evaluation of Two Methods of Grouping," Educational Research Bulletin, 35:213-22, November, 1956.

Another experimental study in which intraclass ability grouping was compared to non-grouping, for instruction in arithmetic, was reported by Wallen and Vowles in 1960.¹⁷ The basis of grouping was the students' arithmetic achievement level. The rotation-group technique was used in order to control the variables. Each round of the experiment consisted of seventy school days. The ungrouped classes all used the same textbooks and were given the same lessons. Each of the four groups of the grouped classes used a textbook appropriate to its level of performance. The teacher met with each group in her class for at least three twenty-minute periods each day. The pupils were encouraged to help each other. No significant differences in achievement were found between the grouped and ungrouped classes.

H. G. Shane named several problems involved in making any type of grouping arrangement successful.¹⁸ Some of those named were:

1. Uneven social and academic profiles of individual children. There is often a variation by over a year from subject to subject.
2. Differences in the philosophy, experience, and competence among teachers.
3. The competence of the administrative leadership.
4. The nature of the curriculum.
5. The nature of the teaching aids and policies governing their use.
6. The size of the school. Some plans need large schools; others need smaller schools.

¹⁷Norman E. Wallen and R. O. Vowles, "The Effect of Intraclass Ability Grouping on Arithmetic Achievement in the Sixth Grade," Journal of Educational Psychology, 51:159-63, June, 1960.

¹⁸H. G. Shane, "Grouping in the Elementary School," Phi Delta Kappan, 41:313-19, April, 1960.

Shane concluded that, "The philosophy and ability of the able teacher are undoubtedly more important than any grouping plan, however ingenuous it may be, with respect to creating a good environment for teaching and learning."

In a critical review by Eckstrom, the point was reiterated that ability grouping is only useful where materials and methods are adapted to the needs of the group. She stated that the evidence slightly favors the use of ability grouping. The greatest relative effectiveness of homogeneous grouping is indicated for dull children, the next greatest for average children and the least benefit and "frequent harm," occur for the bright children.¹⁹ This last statement concurs with the findings of many other studies.

Eckstrom's first point was repeated again by Eash.²⁰ He stated:

Ability grouping in itself does not produce improved achievement. Improved achievement seems to result from manipulation of other factors e.g. curriculum adaptation, teaching methods and materials.

This point seems to be agreed upon by all the writers. However, he added:

Ability grouping may be detrimental to average and low pupils. They suffer from the deprivation of stimulation of brighter children, but bright ones don't suffer from being with duller ones.

This last point was not agreed upon by all writers, particularly in regard to academic achievement.

Wrightstone investigated some of the experiments done in

¹⁹Ruth B. Eckstrom, "Experimental Studies of Homogeneous Grouping: A Critical Review," School Review, 69:216, Summer, 1961.

²⁰Maurice J. Eash, "Grouping: What We Have Learned," Educational Leadership, 18:430, April, 1961.

intraclass and interclass grouping, and found conflicting evidence between the findings.²¹ He also found arguments for and arguments against ability grouping of any kind. He concluded however, that the bulk of the evidence favored intraclass grouping as compared to either interclass or class-as-a-whole instruction.

From a review of this section of the literature, it appears that most, if not all of the writers, agreed that for grouping to be more effective than non-grouping, courses of study, instruction, and materials must be adapted to each group.

A second point that many agreed on, is that grouping appears to be most effective for the slower students, and becomes progressively less effective, with increasing pupil ability. If, however, curriculum were suitably adapted to the needs of each group, grouping might prove equally beneficial to the brighter students.

III. LATER RESEARCH ON INTERCLASS GROUPING

Cochran reported on a flexible grouping program for grade seven, eight and nine.²² Three groups were established for grades seven and eight mathematics and English, and three for grade nine English. Two groups were set up for grades seven and eight science. Pupils were grouped on the basis of I.Q., achievement, and teacher's judgement.

²¹J. Wayne Wrightstone, "Class Organization for Instruction," What Research Says to the Teacher, No. 13, Department of Classroom Teachers and American Educational Research Association (Washington, D.C.: National Education Association, 1957), pp. 1-33.

²²John R. Cochran, "Grouping Students in Junior High School," Educational Leadership, 18:414-19, April, 1961.

Students could be placed in one group level in one subject and a different group level in another subject, and they could be shifted during the year. The curriculum, materials, and teaching procedure were adapted to each group. The control pupils were selected by pairing them with the experimental group.

Only three significant differences were found in the twenty-two comparisons that were made. Cochran concluded, "This data tended to bear out the results of many other studies--that grouping increased achievement, but not to a significant degree."²³

A combined interclass, intraclass type of grouping, referred to as the "Joplin" plan, was reported on by Floyd in 1954.²⁴ This experiment was tried out in Joplin, Missouri, where five teachers of grades four, five and six rearranged their classes for three-quarters of an hour of reading each day. All the pupils in the five classes were given the Iowa reading test and their scores were ranked from highest to lowest. The pupils were then divided into nine equal groups so that four teachers each had two reading groups, and one teacher had one reading group.

No control group was used with which to make comparisons. The pupils were re-tested after four months. It was found that the grade four pupils gained 6.5 months, the grade fives gained 8.7 months and the grade six pupils gained 13.5 months. Because everybody, including

²³Ibid., p. 418.

²⁴Cecil Floyd, "Meeting Children's Reading Needs in the Middle Grades: A Preliminary Report," Elementary School Journal, 15:99-103, October, 1954.

pupils, teachers, and parents, were enthusiastic about the results, it was considered a success.

It is doubtful whether the findings of Floyd's study have any particular statistical value, but the idea for the plan was picked up and tried out, with better controls, by other investigators. Four of these investigations are reviewed briefly here.

Barbe and Waterhouse used a similar type of organization in 1955 in Tennessee.²⁵ Six teachers and one-hundred and eighty pupils of grades four, five and six, were involved over a period of six months. They concluded that:

1. The data collected clearly indicates that a great deal of progress can be made when children are grouped for reading instruction.

2. The needs of children can more nearly be met when they are grouped according to reading ability.

A study by Hart, using the Joplin Plan type of organization, also had results favorable to the experimental groups.²⁶ Thirteen ability groups were distributed among nine teachers of grades four, five and six. The experimental classes exchanged rooms for a block of eighty-five minutes each day for language arts instruction. Control classes, "equal in every way," were set up for comparative purposes. The first experimental period ran from September 1955 to September 1956. The experiment was repeated from September 1956 to September 1957.

²⁵Walter B. Barbe and Tina S. Waterhouse, "An Experimental Program in Reading," Elementary English, 33:102-04, February, 1956.

²⁶R. H. Hart, "Effectiveness of an Approach to the Problem of Varying Abilities in Teaching Reading," Journal of Educational Research, 52:228-31, February, 1959.

In the first year, a "t" score of 7.72 showed a highly significant difference in favor of the grouped classes. In the second year, a "t" score of 6.81 was again highly significant in favor of the grouped classes.

In 1960, Morgan and Stucker used a similar type of arrangement to observe the effects on ninety matched pairs of fifth and sixth grade pupils.²⁷ A one-year instructional program was carried out by eight teachers who were randomly assigned to the experimental and control groups. The results tended to favor the experimental, Joplin Plan, groups.

A study by Koontz, in 1961, on a Joplin Plan type of arrangement, had findings contrary to the others.²⁸ His investigation appears to have been the best controlled of the five. Pupils in grade four were grouped in arithmetic, language and reading on the basis of their scores achieved on the Iowa Tests of Basic Skills. The pupils changed to different classes for study in each of arithmetic, reading, and language. Control pupils were matched with experimental pupils in each of the three subjects.

Koontz found that this organization for grouping had no significant effect on the pupils' achievement.

The majority of the research on grouping has been concerned with the academic achievement of pupils. Some educators however, have been

²⁷ Elmer F. Morgan, Jr., and Gerald R. Stucker, "The Joplin Plan of Reading vs. A Traditional Method," Journal of Educational Psychology, 51:69-73, April, 1960.

²⁸ W. F. Koontz, "Study of Achievement as a Function of Homogeneous Grouping," Journal of Experimental Education, 30:249-53, December, 1961.

concerned with the social effects. In order to point out the dangers associated with ability grouping, particularly the social dangers, Baker reviewed some of the current British and American educators' thoughts on grouping.²⁹ In the summation of his article he stated:

. . .General ability grouping by classes, streams or schools involves difficulties and dangers that demand our concern. The difficulties relate mainly to uncertainties about the nature and growth of intelligence and to problems involved in its measurement. The dangers relate mainly to the nature and extent of educational opportunity for students both as individuals and as developing members of society. I have certainly tried to indicate that where some measure of general ability grouping seems feasible and desirable or necessary, there should remain in the school program a sphere of common activity appropriate to the development of democratic citizenship.

Other educators have also been concerned with the effects of ability grouping on social development. One study was set up in an attempt to discover the effects of interclass ability grouping on the social attitudes of pupils.³⁰ It found that homogeneous grouping of this type tended to create a kind of caste system in the school.

Another study, similar to the Luchins', found that the pupils in the low ability groups developed feelings of inferiority.³¹ They suggested that this was one reason why homogeneous grouping by classes was abandoned in the thirties and we should be careful not to repeat the same mistakes in the sixties.

²⁹H. S. Baker, "Second Thoughts on Grouping," A.T.A. Magazine, 39:6-10, 34, March, 1959.

³⁰Abraham S. Luchins and E. H. Luchins, "Children's Attitudes Toward Homogeneous Grouping," The Journal of Genetic Psychology, 72:3-9 March, 1948.

³¹Maxine Mann, "What Does Ability Grouping Do to the Self-Concept?" Childhood Education, 36:357-60, April, 1960.

Livingstone, referring to a different study, came to different conclusions.³² The study referred to, concerned the personal and social adjustment of fifth grade pupils who had passed through a semi-departmentalized organization in the elementary grades. The S.R.A. inventory was administered to 831 of these pupils. Their adjustment ratings were compared to the ratings of the pupils from the self-contained classes who comprised the norm group used in standardizing the test.

He found the 831 pupils to be superior to those in the self-contained classes in each of the five areas of social adjustment measured by the inventory. He concluded that semi-departmentalization does not hinder a pupil's personal and social adjustment.

A review of a portion of the literature on interclass grouping reveals very similar findings to those on intraclass grouping. There are conflicting findings in both. They both indicate the need to fit the curriculum to the level of the group to be successful. If there is any difference in the effects on pupils between interclass and intraclass grouping, it appears to be in the area of social and personal adjustment.

IV. INDIVIDUALIZED VERSUS GROUP READING

No type of grouping has proven to be a panacea to the problem of individual differences. No matter how carefully the pupils are grouped for instruction, there is still a range of achievement within the group. Various attempts have therefore been made to provide instruction on an

³²Hugh A. Livingstone, "Does Departmental Organization Affect Children's Adjustment?" Elementary School Journal, 61:217-20, January, 1961.

individual basis, particularly in reading in the elementary grades. Six articles are reviewed here therefore, dealing with this topic.

Dolch said that individualized reading implies at least five different things:³³

1. The children are at their seats or at reading tables. They are not in a group before the teacher.
2. Each child has a different book that he himself has selected.
3. Each child receives individual help from the teacher or teacher's helper when he needs it.
4. There is always assumed to be a "sharing period" during which the child shares something of what he has learned with part or all of the other children.
5. It is assumed that certain skills have to be taken up with the class as a whole, or in a group session with the part of the class which needs to learn the skill.

He added that part of the individualized method is to have each child come up to the teacher's desk and read part of his story to her, and that part of the method is to have the teacher moving around.

Dolch stated that the idea that any one book can be suitable for all to appreciate is no longer believed in. Attempts to get a whole group to think and feel about one book at the same time have been a dismal failure.³⁴

Other writers, however, seemed to have other ideas about individualized versus group instruction. Some found that grouping is superior to individualized instruction, and some feel that the best approach is

³³E. W. Dolch, "Individualized Reading vs. Group Reading," Elementary English, 38:566-75, December, 1961.

³⁴Ibid., p. 572.

to use both grouping and individualized instruction.

Don Parker argued, in opposition to Dolch, that the skills should be taught on an individual basis, and the class should be grouped for the education part of the process, to give the pupils an opportunity to use the skills they have learned to generate and apply knowledge in a life-like situation.³⁵

Sartain conducted an experiment to determine whether grade two pupils would make greater gains in reading skills when taught for three months by the method of individualized self-selection, or when taught for an equivalent period by the method of ability grouping using basic readers plus a variety of supplementary textbooks.³⁶

The samples consisted of ten classes of grade two pupils chosen at random from six public schools in Roseville, Minnesota. Five of the ten classes were randomly chosen to participate in the individual self-selection program. The other five classes were taught in three or four ability groups per room. At the end of fifty-six school days, the pupils were tested and the teachers switched methods for another fifty-six days.

He found that there were no significant differences except in word recognition, where the slow readers in the ability groups, gained significantly more than those on the individualized method.

A study was done by Safford to check the value of the

³⁵Don H. Parker, "When Should I Individualize Instruction?" (Chicago: Science Research Associates, Inc.,) p. 1. (Mimeographed.)

³⁶Harry W. Sartain, "The Roseville Experiment With Individualized Reading," Reading Teacher, 13:277-81, April, 1960.

individualized reading method.³⁷ Seven classes of 183 children in grades three to six were given reading instruction by the individualized method. The California Achievement Test Battery measured the growth of the children over a one-year period. The mean growth increments of the seven classes were compared to the national norm of 1.0 year and with the district norm of 1.25 years of growth in reading.

Safford discovered that none of the seven classes came even close to the national norm of 1.0 year, let alone the district's higher norm of 1.25 years. The highest gain was .76 years, and the lowest .16 years. He concluded that for the majority of the pupils, individualized reading techniques resulted in lower gains in reading achievement over a period of a year than other current methods in the same district.

In Witty's article, the main idea was that both individualized reading and ability grouping have something to offer.³⁸ He stated that, although the studies emphasize the value of individualized reading, other studies demonstrate unmistakable values for group procedures. He added, "It seems that a defensible program in reading will combine the best features of both individualized and group instruction in reading."³⁹

The same point was made by Artley.⁴⁰ He concluded:

³⁷Alton L. Safford, "Evaluation of an Individualized Reading Program," Reading Teacher, 13:266-70, April, 1960.

³⁸Paul Witty, "Individualized Reading--A summary and Evaluation," Elementary English, 36:401-12, 450, October, 1959.

³⁹Ibid., p. 410.

⁴⁰A. S. Artley, "Eclectic Approach to Reading," Elementary English, 38:320-7, May, 1961.

On the basis of the judgement and opinion of qualified leaders in the field. . .there seems to be no valid reason for concluding that one must make a choice between individualized reading and a group approach using basal materials. Rather, the wise procedure would be to combine and adapt the best features of each into a pattern that adequately serves the needs of the learner. At times this will involve group procedures, at other times, individual.

Thus, of the six articles reviewed here, only one condemns ability grouping outright, another condemns individualized instruction outright, and four suggest a combination of both methods. Two of the six suggest that ability grouping, combined with the individual method where suitable, should be the core of organization.

V. ACCELERATION AND ENRICHMENT

One alternative to grouping for instruction would be the use of acceleration and enrichment to look after the needs of the very bright pupils. A review of the literature indicates that such an alternative is worth considering.

Hall, after reviewing the literature on enrichment in 1958, arrived at several conclusions.⁴¹ He noted that to provide enrichment properly, adequate equipment, reduced class loads and ample time for teacher planning are required. He found also that enrichment and acceleration should compliment each other in educating the gifted child. There was no unanimity in the studies regarding what should be done for these children.

In a review of the literature on acceleration, Hall found

⁴¹L. Hall, "A Review of Enrichment Procedures for Gifted Pupils," Alberta Journal of Educational Research, 4:117, June, 1958.

that:⁴²

1. Research shows that gifted children who are not accelerated are more likely to develop behavior and personality problems.

2. There are dangers involved in skipping--danger to health, gaps in knowledge, and difficulty is often encountered in later grades.

3. Acceleration should depend on the physical, emotional, mental and social development of the individual child.

4. Terman states that children with 135 I.Q. or higher should be accelerated at least one year. He found little evidence of maladjustment from acceleration, and a follow-up study of accelerated pupils showed that they were doing well at University and many held a large proportion of executive type positions.

5. Another study supported Terman's and found both social and academic advantages accruing to the accelerated pupils.

Hall noted that there was disagreement among authorities as to which was the best for gifted students--acceleration or enrichment or both. Probably both are needed and probably the amount of acceleration depends on the individual student.

Baldauf conducted an experiment to discover whether separating the gifted and enriching their program of studies would increase their academic achievement.⁴³ His controlled experiment showed that achievement is increased. He felt however, that some factors, such as unconscious enrichment for the control groups, were not properly controlled, and therefore the findings may not have been valid. According to Baldauf, however, his findings substantiated those of many similar studies.

⁴²L. Hall, "A Review of Acceleration Procedures for Regular and Special Classes," Alberta Journal of Educational Research, 4:50-57, March, 1958,

⁴³Robert J. Baldauf, "A Comparison of the Extent of Educational Growth of Mentally Advanced Pupils in the Cedar Rapids Experiment," Journal of Educational Research, 52:181-83, January, 1959.

An experiment on enrichment was reported by West and Sievers in 1960.⁴⁴ In an attempt to give advanced study to superior students in arithmetic and language without fostering undesirable attitudes, one school system separated their superior students for half of each day for enriched work and returned them to the heterogeneous regular classes for the other half day. The superior achievers gained over one year more than the regular classes on the Iowa Basic Study Skills Test.

Furthermore, the experimental groups had fewer problems in the area of human relationships and maintained normal relationships with the other students.

Dockerell, after reviewing forty-seven studies and reviews, stated:⁴⁵

The evidence of the experimental studies is quite clear. Merely separating youngsters into different rooms on the basis of intelligence or achievement is valueless. Much more than this is required for grouping to have any effect on achievement. It seems clear that very carefully prepared programs of enrichment are necessary, and that their preparation is difficult.

Wilson reviewed part of the literature and concluded that in general, acceleration is beneficial and that more harm than good can come of denying gifted pupils acceleration appropriate to their degree of giftedness.⁴⁶

Paul Witty reviewed several studies done on providing for gifted

⁴⁴J. West and C. Sievers, "Experiment in Cross-Grouping," Journal of Educational Research, 54:70-2, October, 1960.

⁴⁵W. B. Dockerell, "Special Education for Gifted Children," Canadian Education and Research Digest, 2:37-44, March, 1962.

⁴⁶F. T. Wilson, "Evidence about Acceleration of Gifted Youth," School and Society, 73:409-10, June, 1951.

pupils.⁴⁷ He found that there was disagreement among educators concerning grouping. Some support separation for special instruction only, and some favor enrichment within regular classrooms.

From this brief review on acceleration and enrichment it appears that the majority of the writers support some form of enrichment, or acceleration, or both. They are not agreed on what form these should take. It may be necessary to search for some other method which will look after the needs of the extremely high ability pupils.

Further reference is made to acceleration and enrichment in Chapter V under the "Implications of the Study" section.

VI. SUMMARY

Many of the studies done on grouping took place in the 1920's and 1930's. Most of these made very little, if any, attempt to control the variables. Of those in which an attempt was made to control the variables, some reported that part of this control was done by keeping course content, instructional methods and materials identical for all the groups regardless of ability level. This, of course, would defeat the purpose of grouping, which is to make it possible to deal more efficiently with individual differences. On the other hand, where an attempt was made to provide different course content and materials, often a control group was missing.

In the later studies, most of which took place after 1950, there

⁴⁷Paul Witty, "Nature and Extent of Educational Provisions for the Gifted Pupil," Educational Administration and Supervision, 37:65-79, February, 1951.

was usually much better control of the variables and more sophisticated statistical treatment of the data, but there was such a wide variety in the type of grouping, that it is nearly impossible to find two studies that are alike and have the same results. Of those which appear to be quite similar, there is often disagreement in the findings.

Many of the writers warn of the dangers to the social and personal development of the children involved in grouping, but other studies have shown that certain types of grouping can be organized without harm to the child's social and personal growth.

Organizing classes for individual reading instruction appears to be of most benefit when combined with a core of basic readers and with grouping for instruction.

Because many of the studies have found grouping in its various forms was of benefit, it would be unwise to abandon attempts at providing for individual differences through grouping.

CHAPTER III

THE DESIGN OF THE STUDY

A comparison was made of the reading growth of 365 grade seven students who were grouped within their classes for reading, and the reading growth of 370 grade seven students who were not grouped for reading. The testing period covered seven months. Each of the three reading groups in the grouped classes was given different reading material from the textbook as set out in the Programme of Studies for Elementary Schools for the Province of British Columbia.

During the week of August 20, 1963, the investigator made a trip to Grande Prairie and Dawson Creek to visit all of the superintendents, four of the principals, and the one supervisor concerned. The visit was made in order to explain the purpose of the study, to ask the superintendents' permission to carry out the testing, and to enlist their aid in supervising the testing. It was also necessary to make certain that there would be no grouping in the Alberta schools, and that there would be grouping for reading in the British Columbia schools.

The two school programs. An attempt has been made to specify those variables, such as socio-economic status of the communities and teacher qualifications and competence, which have been assumed to be the same for both groups. Support for the underlying assumption of this study, that the crucial variable of the two groups was method of instruction, (grouping versus non-grouping) has been drawn from statements made

by responsible school personnel.

The following are excerpts from letters received from two of the superintendents and one of the principals in the Grande Prairie schools.

H. Toews wrote:

. . . Please have my assurance that the ten schools and the grade seven classes contained therein do not use the practice of grouping for Literature instruction. Even though grouping may have certain merits, the grade seven teachers in this area do not seem to have had the urge or have found it necessary to use grouping procedures.¹

H. C. Rhodes, superintendent, wrote, "None of the [three] schools grouped the grade seven pupils."²

P. J. Husby, principal, wrote:

Please be advised that the grouping of Grade VII students at this school is done on the basis of their stated interest in the exploratory subjects. There is no grouping of Grade VII students in terms of achievement in the academic subjects.³

These three letters cover all the grade seven students that took part in the study from the Grande Prairie area.

During the visit to Dawson Creek, the superintendent, the supervisor of elementary instruction, and the principals with whom the study was discussed made it clear that the grade seven classes would be grouped for reading, and that differentiation of instruction would take place according to the directions given in the programme of studies. R. I. McLoughlin, Curriculum Consultant for the Department of Education,

¹"Personal Correspondence of the Investigator, letter from H. Toews, Superintendent of Schools, County of Grande Prairie, No. 1, Feb. 18, 1965."

²"Personal correspondence of the Investigator, letter from H. C. Rhodes, Superintendent, School Division No. 54, and Grande Prairie Inspectorate, Feb. 17, 1965."

³"Personal Correspondence of the Investigator, letter from P. J. Husby, Principal, Montrose Junior High School, Grande Prairie, Feb. 16, 1965."

Victoria, B. C. (who was then the Supervisory of Elementary Instruction for the B. C. schools which took part in the study) wrote:

. . .I can state that in my capacity as Supervisory of Elementary Instruction in the Peace River South School District during the 1963-1964 school year, I visited the grade seven classrooms regularly and found that the teachers were attempting to differentiate the reading instruction of their reading groups, as per the instructions outlined in the Programme of Studies for Elementary Schools, Supplement, Grade VII.⁴

Specifically, this meant that the top group would be given those reading selections marked with an "A," the middle group would be given instruction in those marked with a "B," and the low group would be given instruction in those reading selections marked with a "C," as set out in the Programme of Studies. Furthermore, the low reading groups would be given drill in the fundamental reading skills, and their reading difficulties would be analyzed so that remedial and developmental practice could be given "to raise their reading level."⁵ The top reading groups would be given supplementary activities such as the writing of poetry, drama, and creative prose, with such student projects as class newspapers, solutions to unsolved mysteries, fictional interviews with great men of the past, and imaginary diaries of historical or literary figures. In addition they would be given free access to the library, an opportunity to help plan their work, and encouragement to gather data, pictures, and instructional material related to the course.⁶ The details

⁴"Personal Correspondence of the Investigator, letter from R. I. McLoughlin, Curriculum Consultant, Department of Education, Victoria, B. C., Feb. 17, 1965.

⁵Division of Curriculum, Department of Education, Province of British Columbia, Programme of Studies for Elementary Schools, Supplement, Grade VII. (Victoria: Queen's Printer, 1961), pp. 8-34.

⁶Ibid., p. 10.

of the differentiation of instruction would vary from school to school depending upon the materials available and the ability of the individual teacher.

Principals in the Dawson Creek area compiled class lists indicating in which of the three groups the pupils had been placed.⁷ The results below show the group classifications for each school.

<u>School</u>	<u>Top Group(A)</u>	<u>Middle Group(B)</u>	<u>Low Group(C)</u>	<u>Totals</u>
A	22	31	26	79
B	17	29	24	70
C	15	20	5	40
D	21	24	18	63
E	8	8	9	25
F	10	10	8	28
G	7	9	7	23
H	<u>13</u>	<u>16</u>	<u>8</u>	<u>37</u>
Totals	113	147	105	365

Twenty-eight teachers were involved in the study with thirteen teaching the grouped classes and fifteen teaching the non-grouped classes.

I. SELECTION OF SAMPLE AND TESTING PROCEDURE

The mental ability test was given in October so that a comparison could be made of the distributions of the mental ability of the two sets of pupils. No student had an I.Q. which was extremely different from the others. No significant differences in the variances or in the means of mental ability were found. Therefore all the students could be included in the sample as far as mental ability was concerned.

⁷"Personal Correspondence of the Investigator, class lists from M. F. Ryan, H. L. Rodger, B. Ronaghan, K. Smith, J. E. Bastin, D. A. Pritchard, J. G. Exner, and B. Baxter, November, 1964."

The silent reading test was used to assign pupils to one of the three reading groups within the grouped classes. It was also used as a pre-test. It was given again on April 30, 1964, as a post-test. At the time of the pre-test, there were 888 pupils in the study, with 449 grouped pupils and 439 ungrouped pupils. By the time of the post-test, through absenteeism, or transfers, or non-completion of tests, a total of 124 pupils' scores were removed. This left 376 grouped pupils and 388 ungrouped pupils. The distributions of reading scores on the pre-test of these remaining students were compared. Using the "F" test and the "t" test, an "F" ratio of 1.11 was found.⁸ This is not significant at the .05 level, so that homogeneity of variance was assumed. However, a "t" ratio of 4.03 was found. This is significant at the .05 level. Therefore it was assumed that there was a difference in the means.

Since it was desirable that there be no significant differences in the means of the reading scores on the pre-test, eleven more grouped students' scores were removed and eighteen ungrouped students' scores were removed. This was done by comparing the graphs of the two distributions of reading scores on the pre-test, and removing scores from the areas on the graphs where large differences were indicated. The same students' scores were also removed from the post-test scores. The results are as shown below, and the graphs are shown in Figure 1, p. 45. The hatched portions of the graph are the areas from which the twenty-nine pupils were removed.

⁸George A. Ferguson, Statistical Analysis in Psychology and Education (Toronto: McGraw-Hill Book Company, Inc., 1959), pp. 131-146.

TABLE I
FREQUENCY DISTRIBUTIONS OF READING SCORES ON PRE-TEST

Reading Score	Grouped Students	Ungrouped Students
186 - 188	1	1
183 - 185	4	3
180 - 182	6	2
177 - 179	17	7
174 - 176	19	21
171 - 173	27	19
168 - 170	29	19
165 - 167	16	22
162 - 164	26	28
159 - 161	35	36
156 - 158	35	36
153 - 155	37	29
150 - 152	30	38
147 - 149	21	35
144 - 146	23	31
141 - 143	17	23
138 - 140	14	22
135 - 137	10	8
132 - 134	8	4
129 - 131	1	4
126 - 128	0	0
N =	376 ^a	388 ^b

^aEleven students' scores removed later.

^bEighteen students' scores removed later.



FIGURE 1

SCORES OF GROUPED AND UNGROUPED STUDENTS ON PRE-TEST AFTER
REMOVAL OF STUDENTS WHO MISSED POST-TEST^a

^a Twenty-nine students were later removed from the hatched area.

	<u>Pre- test</u>	<u>Absentees, and others</u>	<u>Post- test</u>	<u>Removed</u>	<u>Total Left</u>
Grouped	449	73	376	11	365
Ungrouped	<u>439</u>	<u>51</u>	<u>388</u>	<u>18</u>	<u>370</u>
Totals	888	124	764	29	735

With these final twenty-nine scores removed, the reading score distributions were again compared using the "F" test and the "t" test. An "F" ratio of 1.11 was found. This is not significant at the .05 level, so it was concluded that there was still homogeneity of variance. The "t" ratio was found to be 1.63, and there was now no significant difference in the means at the .05 level.

With the overall reading score distributions on the pre-test now made comparable, the following additional comparisons were made on the results of the pre-test and the post-test, in order to test the null hypotheses. The "F" test and the "t" test were used in each case to test for the significance of difference.

The mean and variance of the grouped students' scores were compared to the mean and variance of the ungrouped students' scores on the overall reading test.

The means and variances of the above-average students' scores in both samples were compared on the overall reading test.

The means and variances of the below-average students' scores in both samples were compared on the overall reading test.

The means and variances of the students' scores in both samples were compared on the reading sub-test "comprehension."

The means and variances of the above-average students' scores in

both samples were compared on the reading sub-test "comprehension."

The means and variances of the students' scores in both samples were compared on the reading sub-test "word meaning."

The means and variances of the below-average students' scores in both samples were compared on the reading sub-test "word meaning."

II. A COMPARISON OF GROUPED AND UNGROUPED STUDENTS IN RELATION TO BACKGROUND AND INTELLIGENCE

Background

It has been assumed that the pupils in both samples had a similar background in reading in the elementary grades.⁹

The cities of Dawson Creek and Grande Prairie are within ninety miles of each other and both are the centres of farming communities. The grouped classes were drawn from Dawson Creek and its surrounding area for the study, and the ungrouped classes from Grande Prairie and its surrounding area. Therefore the school populations were considered comparable in socio-economic and cultural background.¹⁰ Furthermore, the grouped classes received a different treatment, namely, three groups per class and differentiated instruction while the ungrouped classes received the typical treatment, namely, no grouping and no differentiation of instruction.¹¹

Intelligence

The Otis Quick Scoring Mental Ability Test, Gamma, Form Em, was

⁹See pp. 7,8 above. ¹⁰See p. 6, above. ¹¹See pp. 40-42, above.

given on the same day that the reading pre-test was given. The mean for the grouped students was 105.14 and for the ungrouped 105.97. When these means were tested for significance of difference, a "t" value of 1.04 was obtained. This difference is not significant at the .05 level.

The variances of the two samples were also compared. The grouped pupils had a variance of 107.85 and the ungrouped 123.99. From these variances an "F" ratio of 1.15 was obtained which is not significant at the .10 level.

The mean of the above-average students in the grouped sample was 116.55 and the mean for the ungrouped was 118.32. The difference of 1.77 gave a "t" value of 1.26 which is not statistically significant.

The means of the below-average students in each sample were then compared. A "t" value of .46 again indicated no significant difference.

The variances of the above-average pupils produced an "F" ratio of 1.06. The variances of the below-average pupils produced an "F" ratio of 1.44. Neither of these ratios indicate a significant difference.

Since the means and variances of the intelligence quotients were found to be not significantly different, it was assumed that any differences found in reading achievement could not be attributed to intelligence.

III. THE TESTING INSTRUMENTS

One writer, in reviewing the Otis Quick-Scoring Mental Ability Tests, said:

The Otis Quick Scoring Mental Ability Tests, as the title implies, do furnish a short and easily scored indicator of scholastic aptitude. Such a measure, if interpreted with care, can be

TABLE II
FREQUENCY DISTRIBUTION OF I.Q.'S OF EACH SAMPLE

I.Q. Rating	Grouped Students	Ungrouped Students
140 - 144	1	1
135 - 139	0	3
130 - 134	3	5
125 - 129	14	10
120 - 124	14	19
115 - 119	35	42
110 - 114	48	55
105 - 109	65	61
100 - 104	70	77
95 - 99	69	42
90 - 94	21	29
85 - 89	22	17
80 - 84	3	8
75 - 79	0	1
N =	365	370
Mean =	105.14	105.97
S. D. =	10.39	11.14
s^2 =	107.85	123.99
"t" =		1.05
"F" =		1.15

TABLE III

I.Q. DISTRIBUTIONS OF ABOVE-AVERAGE AND BELOW-AVERAGE STUDENTS

I.Q. Rating	Above-Average		Below-Average	
	Grouped	Ungrouped	Grouped	Ungrouped
140 - 144	1	1		
135 - 139	0	3		
130 - 134	2	5		
125 - 129	13	7		
120 - 124	10	11		
115 - 119	22	20	0	0
110 - 114	15	16	2	3
105 - 109	8	5	3	5
100 - 104	5	4	12	15
95 - 99	2	0	20	13
90 - 94	0	0	10	11
85 - 89			12	14
80 - 84			3	7
75 - 79			0	1
N =	78	72	62	69
Mean =	116.55	118.32	95.47	94.83
S.D. =	8.55	8.80	7.00	8.40
S ² =	73.10	77.44	49.00	70.56
"t" =		1.26		.46
"F" =		1.06		1.44

useful to both teacher and counselor by revealing within fairly broad limits of accuracy the probable level of academic achievement for a majority of pupils.¹²

Another review from the same source said that if it is required to segregate pupils into homogeneous groups for teaching, administrative, or research purposes, this test affords a satisfactory rough classification.¹³

The Iowa Tests of Basic Skills are recommended by both the Alberta and B. C. courses of studies. These tests have also been reviewed in The Fifth Mental Measurements Yearbook. One review stated:

The real strength of the tests is in their curricular validity, careful construction, provision of adequate norms based on a national sample, and high reliabilities. The manuals provide the teacher with excellent help in using test results to improve instruction. These advantages outweigh the disadvantages of length and time necessary for administration. The tests, for their purposes, are among the best available at this time.¹⁴

Another author from the yearbook noted that technically the tests are beyond criticism.¹⁵

A third writer, in the same yearbook said:

No battery of achievement tests intended for civilian uses has been constructed with greater technical sophistication, greater adequacy of statistical base, and greater use of previous research.¹⁶

Thus both tests appear to be quite satisfactory for the purpose of this study.

¹²Oscar Krisen Buros (ed.), The Fifth Mental Measurements Yearbook (Highland Park, New Jersey: The Gryphon Press, 1959), p. 499.

¹³Ibid., p. 499. ¹⁴Ibid., p. 33.

¹⁵Ibid., p. 35. ¹⁶Ibid., p. 36.

IV. STATISTICAL TREATMENT

In testing the hypotheses, listed in Chapter I, the difference between the means for independent samples was tested.¹⁷ Since the form of the "t" test for independent samples required homogeneity of variance, and since it was part of the study to determine whether there was any significant difference in range of achievement, the "F" test was applied first to the variances.¹⁸

After the mean, standard deviation and variance were calculated for each of a pair of observations, the variance was tested with the formula:

$$F = \frac{S_1^2}{S_2^2}$$

The appropriate "t" ratio was then obtained in three steps:

1. An unbiased estimate of the population variance was determined by the formula:

$$S^2 = \frac{(N_1 - 1)S_1^2 + (N_2 - 1)S_2^2}{N_1 + N_2 - 2}$$

2. The unbiased variance estimate S^2 was used to obtain an estimate of the standard error.

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S^2}{N_1} + \frac{S^2}{N_2}}$$

3. The differences between the means was then divided by the estimate of the standard error.

¹⁷Ferguson, op. cit., p. 137.

¹⁸Ibid., p. 141; Helen M. Walker and Joseph Lev, Statistical Inference (New York: Holt, Rinehart and Winston, Inc., 1953), p. 156.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

The number of degrees of freedom used in evaluating "t" was $N_1 + N_2 - 2$. Two degrees of freedom are lost because deviations are taken separately about the means of the two samples.

The statistical treatment outlined above was used in the analysis of data which follows in Chapter IV.

CHAPTER IV

ANALYSIS OF DATA

The results of all tests were recorded on appropriate sheets. The raw scores of the reading tests were converted to standard scores using the scales given at the bottom of each page of the Iowa Silent Reading Test. Each pupil's sub-test standard scores were recorded and the mean standard scores were then calculated for each student and recorded. This was done for each set of scores of the post-test and the pre-test.

The means, standard deviations, and variances were then calculated for each set of scores for which comparisons were made. The variances were tested with the conventional "F" test, using the accepted ten per cent level of confidence.

The significance of the difference between two means was then tested by the conventional "t" test, using the accepted one per cent and five per cent levels. Analysis of the data was made from these calculations.¹

I. COMPARISON OF OVERALL GROWTH ON TOTAL SCORES

The hypothesis that there would be no significant difference in the overall growth of reading achievement between the grouped and ungrouped students was tested by comparing the total scores of each set

¹George A. Ferguson, Statistical Analysis in Psychology and Education (Toronto: McGraw-Hill Book Company, Inc., 1959), pp. 131-146.

of students on the pre-test and again on the post-test.

The mean score obtained by the grouped classes on the pre-test was 157.82 compared with 156.37 by the ungrouped classes. As the "t" ratio was 1.63, it was concluded that no significant difference existed in pre-test scores at the .05 level.

On the post-test, the grouped and ungrouped classes showed mean scores of 165.47 and 164.78 respectively. Again, with "t" equal to .74, the difference was not significant.

The "F" test produced values of 1.11 and 1.03 on the pre-test and post-test respectively, neither of which is significant at the .10 level.

These pre-test and post-test scores are shown in Table IV, and pictured graphically in Figure 2.

II. COMPARISON OF ABOVE-AVERAGE STUDENTS ON TOTAL SCORE

The hypothesis that there would be no significant difference in the growth of achievement between the above-average reading groups, was tested by comparing the total scores of each set of above-average students on the pre-test and again on the post-test.

Seventy-eight of the grouped pupils scored one standard deviation above the mean of the pre-test. Seventy-two of the ungrouped pupils scored one standard deviation above the mean on the pre-test. These students were called "above-average." The above-average students in the grouped classes had a mean score of 175.19. The ungrouped, above-average students' mean score was 173.63. This difference produced a "t" value of 2.31, which is significant at the .05 level.

TABLE IV
TOTAL READING SCORES ON PRE-TEST AND POST-TEST BY
STUDENTS IN BOTH SAMPLES

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
189 - 191	0	0	2	1
186 - 188	1	1	4	10
183 - 185	4	3	14	5
180 - 182	6	2	19	20
177 - 179	17	7	34	24
174 - 176	19	21	25	29
171 - 173	22	19	30	36
168 - 170	23	19	33	37
165 - 167	16	22	36	42
162 - 164	26	28	40	32
159 - 161	35	36	26	23
156 - 158	35	36	28	20
153 - 155	37	29	24	31
150 - 152	30	34	14	20
147 - 149	21	31	18	21
144 - 146	23	27	8	7
141 - 143	17	21	3	6
138 - 140	14	18	4	0
135 - 137	10	8	1	3
132 - 134	8	4	1	1
129 - 131	1	4	1	1
126 - 128	0	0	0	1
N =	365	370	365	370
Mean =	157.82	156.37	165.41	164.78
S.D. =	12.39	11.78	11.39	11.58
S ² =	153.44	138.65	129.75	134.09
"t" =		1.63		.74
"F" =		1.11		1.03



FIGURE 2

COMPARISON OF OVERALL STUDENT SCORES ON PRE-TEST AND POST-TEST

Pre-test		Post-test
Grouped Students	—	Grouped Students
Ungrouped Students	- - -	Ungrouped Students

On the post-test however, the seventy-eight above-average, grouped students, had a mean score of 178.11 compared to a mean of 178.67 for the seventy-two ungrouped students. Thus the difference now favored the ungrouped classes, although a "t" value of .58 shows that the difference was not significant. The results are shown in Table V.

When checked for the difference in range of achievement on the pre-test, it was found that the grouped classes had a variance of 17.36 and the ungrouped had a variance of 17.21. The "F" value of 1.01 showed no significant difference. On the post-test, the grouped and ungrouped classes had variances of 38.75 and 29.32 respectively. An "F" ratio of 1.32 again showed no significant difference.

It was noted that five of the grouped and one of the ungrouped students scored less on the post-test than on the pre-test.

III. COMPARISON OF BELOW-AVERAGE STUDENTS ON TOTAL SCORE

The hypothesis that there would be no significant difference in the growth of reading achievement between the below-average reading groups, was tested by comparing the total scores of both sets of below-average students on the pre-test and again on the post-test.

Pupils who scored one standard deviation below the mean on the pre-test were called "below-average." Sixty-two of the grouped pupils and sixty-nine of the ungrouped fell into this category. The mean score of the below-average grouped pupils was 139.58 and that of the below-average ungrouped was 139.91, on the pre-test. A "t" value of .47 showed no significant difference. On the post-test, the sixty-two

TABLE V

TOTAL READING SCORES ON PRE-TEST AND POST-TEST BY ABOVE-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
189 - 191	0	0	2	1
186 - 188	1	1	4	10
183 - 185	4	3	13	4
180 - 182	6	2	14	17
177 - 179	17	7	22	17
174 - 176	19	21	8	10
171 - 173	22	19	6	7
168 - 170	9	19	4	5
165 - 167	0	0	3	1
162 - 164			0	0
159 - 161			1	0
156 - 158			1	0
153 - 155			0	0
N =	78	72	78	72
Mean =	175.19	173.63	178.11	178.67
S.D. =	4.17	4.15	6.23	5.42
S ² =	17.36	17.21	38.75	29.32
"t" =	2.31 ^a		.58	
"F" =	1.01		1.32	

^aSignificant at the .05 level.

grouped pupils scored a mean of 150.71 and the ungrouped had a mean of 150.43. A "t" value of .21 showed that there was still no significant difference in means.

The grouped classes had a variance of 16.31 on the pre-test compared with a variance of 17.07 for the ungrouped. This gives an "F" value of 1.05 which is not significant at the .10 level. On the

post-test a variance of 62.68 for the grouped and a variance of 52.36 for the ungrouped produced an "F" ratio of 1.20 which is not significant at the .10 level. Table VI shows the distribution of scores of the below-average students.

TABLE VI

TOTAL READING SCORES ON PRE-TEST AND POST-TEST BY BELOW-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
171 - 173			0	0
168 - 170			0	1
165 - 167			2	1
162 - 164			4	4
159 - 161			7	3
156 - 158			3	4
153 - 155			8	11
150 - 152			10	13
147 - 149	0	0	11	16
144 - 146	12	14	8	7
141 - 143	17	21	2	5
138 - 140	14	18	4	0
135 - 137	10	8	1	2
132 - 134	8	4	1	1
129 - 131	1	4	1	1
126 - 128	0	0	0	0
N =	62	69	62	69
Mean =	139.58	139.91	150.71	150.43
S.D. =	4.17	4.13	7.92	7.24
S^2 =	16.31	17.07	62.68	52.36
"t" =		.47		.21
"F" =		1.05		1.20

IV. COMPARISON OF GAINS ON THE SUB-TEST "COMPREHENSION"

The hypothesis that there would be no significant difference in the growth of the reading skill "comprehension," was tested by comparing the sub-test "comprehension" scores of each set of students on the pre-test and again on the post-test.

When the pupils' scores on the "comprehension" sub-test were tested for significant differences, it was found that the grouped students had a mean of 13.46 on the pre-test. The ungrouped had a mean of 12.75. A "t" value of 2.89 showed that the difference was significant at the .01 level. When the two samples were compared on the post-test for this skill, there was an even greater difference in favor of the grouped students giving a "t" ratio of 5.45 which is significant at the .001 level. In neither case was there any significant difference in the variances as shown by the "F" test. The results are shown in Table VII.

Since the difference for this skill was greater on the post-test than on the pre-test, a further analysis was carried out to determine whether the difference was caused by the above-average, or below-average pupils. A significant difference was found in the means of the below-average students on the post-test. The below-average pupils in the grouped sample obtained a mean score of 13.52 compared to a mean of 11.81 for the ungrouped. A "t" ratio of 4.07 showed that this difference is significant at the .001 level. Thus the difference in the total scores on the sub-test seems to be due mainly to the greater growth of the below-average grouped students, and not by the

TABLE VII

COMPREHENSION SCORES ON PRE-TEST AND POST-TEST BY
ALL STUDENTS IN BOTH SAMPLES

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
20	4	3	24	9
19	16	10	39	35
18	28	16	48	27
17	33	27	54	44
16	35	34	52	42
15	39	40	37	38
14	36	39	36	36
13	40	37	23	40
12	31	38	16	36
11	33	31	15	11
10	20	22	9	18
9	16	21	6	10
8	11	19	4	14
7	11	13	0	4
6	4	8	0	1
5	2	9	0	0
4	3	2	1	2
3	1	0	0	0
2	1	0	0	0
1	1	1	0	0
0	0	0	1	3
N =	365	370	365	370
Mean =	13.46	12.75	15.69	14.38
S.D. =	3.56	3.58	2.98	3.51
S ² =	12.69	12.83	8.89	12.31
"t" =	2.89 ^a		5.45 ^b	
"F" =	1.01		1.40	

^aSignificant at .01 level.

^bSignificant at .001 level.

above-average grouped students.

No other significant differences in means were found on this sub-test and no significant differences were found in the variances.

The results are tabulated in Tables VIII and IX.

TABLE VIII
COMPREHENSION SCORES ON PRE-TEST AND POST-TEST
BY ABOVE-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
20	4	3	16	6
19	11	8	13	20
18	16	8	12	10
17	13	11	15	9
16	10	14	7	11
15	7	9	6	7
14	8	8	5	4
13	4	4	2	3
12	3	3	1	0
11	1	2	1	1
10	1	2	0	0
9	0	0	0	1
8			0	0
N =	78	72	78	72
Mean =	16.48	15.89	17.42	17.07
S.D. =	2.24	2.38	2.16	2.26
S^2 =	5.23	5.68	4.65	5.10
"t" =		1.46		.97
"F" -		1.08		1.09

TABLE IX
COMPREHENSION SCORES ON PRE-TEST AND POST-TEST
BY BELOW-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
20			1	0
19	0	0	0	1
18	0	1	2	2
17	2	0	4	4
16	0	1	10	3
15	2	5	9	3
14	4	2	9	8
13	4	4	4	7
12	6	4	6	10
11	9	9	10	3
10	4	5	3	7
9	4	5	3	7
8	6	10	1	9
7	9	7	0	0
6	3	5	0	0
5	2	6	0	0
4	0	1	1	2
3	1	0	0	0
2	0	0		
1	1	1		
N =	62	69	62	69
Mean =	10.11	9.49	13.52	11.81
S.D. =	3.18	3.31	2.84	3.21
S ² =	10.10	10.95	8.06	10.30
"t" =		1.09		4.07 ^a
"F" =		1.08		1.28

^aSignificant at the .001 level.

V. COMPARISON OF GAINS ON THE SUB-TEST "WORD MEANING"

The hypothesis that there would be no significant difference in the growth of the reading skill "word meaning," was tested by comparing the "word meaning" scores of each set of students on the pre-test and again on the post-test.

An analysis of the results shown in Table X, revealed an "F" ratio of 1.14 and a "t" ratio of .38 on the pre-test, neither of which indicates a significant difference. On the post-test, "F" and "t" were 1.37 and .08 respectively. Thus there was still no significant difference.

A further analysis was done on the differences between the above-average students and the below-average students for "word meaning." On the pre-test, for the above-average students, an "F" ratio of 1.89 was found and a "t" ratio of .17, neither of which was significant. On the post-test, the "F" ratio was 1.01 and the "t" ratio was 1.28. Again, there was no significant difference.

A comparison of the two samples of below-average students, on the pre-test, produced an "F" ratio of 1.01 and a "t" ratio of 1.23. On the post-test, "F" was 1.18 and "t" was 1.74. None of these ratios indicated a significant difference.

The results, therefore, as shown in Tables X, XI and XII, reveal no significant differences either in the variances or in the means of the sub-test "word meaning." Nevertheless, the same trend was indicated here as in the other results analyzed, that is, a tendency for the ungrouped students to gain more than their counterparts.

TABLE X
TOTAL WORD MEANING SCORES ON PRE-TEST AND POST-TEST
BY STUDENTS IN BOTH SAMPLES

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
54 - 56	0	0	0	0
51 - 53	2	1	1	3
48 - 50	2	2	5	11
45 - 47	5	11	18	19
42 - 44	18	19	35	45
39 - 41	35	35	75	50
36 - 38	59	49	49	73
33 - 35	48	56	70	51
30 - 32	63	62	37	31
27 - 29	46	32	30	35
24 - 26	36	47	28	27
21 - 23	30	28	11	17
18 - 20	15	15	4	4
15 - 17	4	9	1	2
12 - 14	1	3	1	1
9 - 11	0	0	0	0
6 - 8	1	0	0	0
3 - 5	0	1	0	0
0 - 2	0	0	0	1
N =	365	370	365	370
Mean =	31.75	31.54	35.23	35.27
S.D. =	7.12	7.61	6.43	7.52
S ² =	50.69	57.91	41.34	56.55
"t" =		.38		.08
"F" =		1.14		1.37

TABLE XI
WORD MEANING SCORES ON PRE-TEST AND POST-TEST
BY ABOVE-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
54 - 56	0	0	0	0
51 - 53	2	1	1	3
48 - 50	2	2	5	9
45 - 47	3	10	13	13
42 - 44	10	11	20	15
39 - 41	20	17	25	19
36 - 38	25	16	7	12
33 - 35	10	11	5	1
30 - 32	5	4	0	0
27 - 29	0	0	0	0
24 - 26	1	0	2	0
21 - 23	0	0	0	0
N =	78	72	78	72
Mean =	38.65	39.61	41.47	42.76
S.D. =	4.71	4.74	4.71	4.35
s^2 =	22.71	22.46	22.18	18.92
"t" =		1.23		1.74
"F" =		1.01		1.18

TABLE XII
WORD MEANING SCORES ON PRE-TEST AND POST-TEST
BY BELOW-AVERAGE PUPILS

Scores	Pre-Test		Post-Test	
	Grouped	Ungrouped	Grouped	Ungrouped
48 - 50			0	0
45 - 47			0	1
42 - 44	0	0	1	0
39 - 41	0	1	3	1
36 - 38	1	1	4	4
33 - 35	0	3	11	10
30 - 32	4	5	8	7
27 - 29	7	4	11	13
24 - 26	12	14	12	17
21 - 23	21	18	9	11
18 - 20	13	12	3	3
15 - 17	4	7	0	1
12 - 14	0	3	0	1
9 - 11	0	1	0	0
6 - 8	0	0		
N =	62	69	62	69
Mean =	23.05	22.90	28.96	27.67
S.D. =	4.23	5.82	5.79	5.76
S^2 =	17.89	33.87	33.52	33.18
"t" =		.17		1.28
"F" =		1.89		1.01

VI. COMPARISON OF VARIANCES

The hypothesis that there would be no significant difference between any of the variances was tested by comparing the variances for each set of data as they arose in conjunction with the "t" tests. In no case was there a significant difference in the variances.

VII. SUMMARY OF RESULTS

Table XIII is a summary of the tests for significance of the mean differences and variances for each statistical test. It also indicates which sample achieved the higher score in each case. A positive mean difference indicates a higher score for the grouped classes. A negative mean difference indicates a higher score for the ungrouped classes. Only four of the statistical tests indicated a significant difference.

TABLE XIII
SUMMARY OF TESTS OF SIGNIFICANCE

		Otis I.Q.	Iowa Total Score	Iowa Comprehensive	Iowa Word Meaning
All Students					
Pre-Test	$\bar{D}$	-	+	+	+
	$\bar{X}$	N.S.D.	N.S.D.	.01	N.S.D.
	S^2	N.S.D.	N.S.D.	N.S.D.	N.S.D.
Post-Test	$\bar{D}$		+	+	-
	$\bar{X}$		N.S.D.	.001	N.S.D.
	S^2		N.S.D.	N.S.D.	N.S.D.
Above-Average Students					
Pre-Test	$\bar{D}$	-	+	+	-
	$\bar{X}$	N.S.D.	.05	N.S.D.	N.S.D.
	S^2	N.S.D.	N.S.D.	N.S.D.	N.S.D.
Post-Test	$\bar{D}$		-	+	-
	$\bar{X}$		N.S.D.	N.S.D.	N.S.D.
	S^2		N.S.D.	N.S.D.	N.S.D.
Below-Average Students					
Pre-Test	$\bar{D}$	+	-	+	+
	$\bar{X}$	N.S.D.	N.S.D.	N.S.D.	N.S.D.
	S^2	N.S.D.	N.S.D.	N.S.D.	N.S.D.
Post-Test	$\bar{D}$		+	+	+
	$\bar{X}$		N.S.D.	.01	N.S.D.
	S^2		N.S.D.	N.S.D.	N.S.D.

⁺Difference in means favors grouped students.

⁻Difference in means favors ungrouped students.

N.S.D. No significant difference.

CHAPTER V

CONCLUSIONS AND IMPLICATIONS

After analyzing the data, certain conclusions and opinions were formed. These are stated as they relate to each of the hypotheses in turn.

HYPOTHESIS I

There is no significant difference in the overall growth of achievement between the grouped and the ungrouped students.

From the statistical evidence found, Hypothesis I was accepted. It is interesting to note, however, that the difference in means between the two samples diminished over the seven-month period. If the treatment being given to the grouped classes, that is grouping and differentiation of instruction, was capable of producing higher academic results than the treatment given the ungrouped classes, then it would be expected that the difference in means should have increased, rather than decreased over this period. Therefore, for the two samples in this study, it appears that the type of grouping and differentiation of instruction used, was no more effective a method of organization than non-grouping and non-differentiation of instruction.

Where the three groups in a grouped class are each given instruction suited to their level of ability, it might be expected that the spread in range of achievement would increase not only as compared to the ungrouped classes, but also as compared to themselves. There was

less difference in variance both as compared to themselves, and as compared to the ungrouped classes, at the end of the seven-month period. Furthermore, the range for both samples remained almost identical.

Since the Dawson Creek schools have only begun to group grade seven for instruction since 1961, these results may be partly explained by the possibility that the teachers of the grouped classes have not yet learned how to make the most effective use of grouping procedures. It is also possible that some of the teachers of the grouped classes are not as "keen" about this type of organization as their Department of Education.

It may also be explained in part by the fact that the grouped students all have the same reading text, and the possibility that the difference in instruction is one of quantity.

HYPOTHESIS II

There is no significant difference in the growth of achievement between the above-average reading groups.

There was a significant difference in means in favor of the grouped classes on the pre-test, but there was no significant difference on the post-test. Therefore no conclusion can be reached regarding this hypothesis. However, there appeared to be a tendency here for the students in the above-average grouped classes to gain relatively less than their counterparts.

HYPOTHESIS III

There is no significant difference in the growth of achievement between the below-average reading groups.

On the basis of the data shown in Table VI, Hypothesis III was accepted. The means and variances are almost identical for both groups on both the pre-test and post-test. It would appear, therefore, that the type of grouping and differentiation of instruction used made no difference to the growth in achievement.

An examination of the ranges in Tables V and VI reveals an interesting fact. The above-average students' scores of both the grouped and ungrouped tended to shift downward toward the mean, while the below-average students of both groups tended to spread upward toward the mean. This would indicate that whichever method of organization was used, there was a tendency for all scores to converge at the mean. It appears that the top students were not being helped to realize their fullest potential.

HYPOTHESIS IV

There are no significant differences in the growth of the reading skill "comprehension."

From the evidence shown in Table VII, page 62, Hypothesis IV is rejected. On the sub-test "comprehension," the total number of grouped students scored significantly higher than the total number of ungrouped on both the pre-test and the post-test. It can only be stated from Table VII that the difference tended to increase in favor of the grouped students. However, it can be seen from Table IX, page 64, that the

difference in growth between the two samples came from the below-average grouped students who scored significantly higher on the post-test. These results may be explained in part by a difference in emphasis, on this particular aspect of the reading program. Possibly too, the teachers of the grouped classes spent more time instructing the below-average students, than they did instructing the above-average.

Tables VIII and IX, pages 63, 64, show that the below-average pupils of both samples have a much greater range in achievement than the above-average pupils of both samples. This could be explained partly by the ceiling on this sub-test which has tended to compress the above-average students into a smaller range and thus the findings are probably partly obscured for the above-average students.

HYPOTHESIS V

There is no significant difference in the growth of the reading skill "word meaning."

On the basis of the evidence, Hypothesis V is accepted.

Although there was no significant difference in means, the ungrouped classes had a slightly lower mean on the pre-test for "word meaning" than the grouped classes. The ungrouped had however, a slightly higher mean on the post-test and therefore the same tendency was shown by this sub-test as that for the total test, namely, a tendency for the ungrouped classes to gain more than the grouped classes. Again, a possible explanation may lie in the difference in emphasis on reading skills by the two school systems. The scores by the students on this sub-test do not appear to have been affected by the "ceiling" on the

sub-test since no student received full marks on either the pre-test or post-test. The range was identical for both groups on both tests. The variance for the grouped classes was reduced on the post-test, showing the same tendency for the students to converge at the mean, rather than to spread upwards.

HYPOTHESIS VI

There is no significant difference between any of the variances which are compared.

On the basis of the evidence, Hypothesis VI was accepted. When all the variances were examined, there appeared to be a tendency for them to diminish. This is noticeable in Tables IV, VII, and X. In other words, for both grouped and ungrouped students, there is less spread away from the mean near the end of Grade Seven than there was at the beginning. But in Tables V and VI the variances increased. This was due to the above-average students' scores moving downward toward the mean, and the below-average students' scores moving upward toward the mean. This tendency is hidden in Tables IV, VII, and X because the scores of the above-average and below-average students were distributed among the average scores.

It would appear, therefore, that the instruction given in both school systems was designed mainly for the average, and possibly for the below-average student. The above-average student, although he was placed in a different group in the grouped classes, may not have been given instruction and material well enough suited to his needs.

SUMMARY OF CONCLUSIONS

From Table XIII, page 70, it can be seen that only four significant differences were found in means. There was an overall tendency for no difference to exist in the means on either the post-test or pre-test. However, except for the sub-test of "comprehension," any differences that did exist, showed a tendency for the average, and above-average grouped students to gain less than their counterparts. On the other hand, below-average students tended to gain more when grouped.

Although no significant differences in variances were found, there appeared to be a tendency for the majority of the scores in both samples to move toward the mean. This is possibly due to the instruction, in both school systems, being best suited to the level of the average and below-average students. The above-average grouped students, although they were grouped for instruction, do not appear to have had sufficient differentiation of instruction and materials. It is also possible that they did not receive as much instructional time as the lower groups. Only a few of the below-average students, at the extreme low end of both samples, did not gain over the seven-month period.

The below-average grouped students tended to gain more in relation to the other groups on the total test. They gained significantly more on the "comprehension" sub-test, and showed a strong tendency to gain more on the "word meaning" sub-test. The above-average grouped students lost significantly more in relation to the other groups on the total test, and tended to lose more on both sub-tests. Thus it may be concluded that grouping proved partly beneficial to the below-average

students, but appeared to be of no benefit to the above-average students. Furthermore, those students at both extremes gained little or nothing, no matter which method of organization was used.

IMPLICATIONS OF THE STUDY

In a school system where grouping within the classrooms is mandatory, and the teachers are directed to differentiate the instruction within these classrooms, one might speculate that all of the students should benefit from such treatment when it is given. Most of the above-average students in the present study did not appear to benefit. The results of this study are not too surprising since research indicates that grouping of itself does not guarantee superior academic achievement. Research does indicate that where grouping has produced gains in academic achievement, this grouping has been accompanied by such factors as curriculum adaptation and differentiated teaching methods.

Teachers and administrators would be well advised to see that better use is made of the materials which are available, and attempt to make a better adjustment of instruction to the needs of their groups, if grouping, as a method of organization for instruction, is to become more effective than it is at present.

Since the present study found that the bright students gained relatively less, on the whole, than the other students, a brief review of the literature on acceleration and enrichment was prepared for Chapter II. As indicated in that section, some form of acceleration or enrichment or both would appear to be an acceptable alternative to grouping

as far as the gifted student is concerned. The implication would therefore be that educational administrators should consider such an alternative wherever they attempt to meet the needs of their gifted pupils.

RECOMMENDATIONS FOR FURTHER STUDY

Although much research has been done on grouping, in an attempt to solve the problem of looking after individual differences, many of the problems associated with grouping still have not been solved. The following studies are therefore presented for consideration:

1. A study on the effect of grouping on personal and social development of students.
2. Studies where methods and materials are developed with the specific needs of each of the groups in mind and an evaluation of these programs.
3. Studies on various methods of grouping and their effectiveness at the senior high school level.
4. Studies of the effects of grouping in the primary grades.

BIBLIOGRAPHY

BIBLIOGRAPHY

A. BOOKS

Ferguson, George A. Statistical Analysis in Psychology and Education. Toronto: McGraw-Hill Book Company, Inc., 1959.

Walker, Helen M., and Joseph Lev. Statistical Inference. New York: Holt, Rinehart and Winston, Inc., 1953.

B. PUBLICATIONS OF THE GOVERNMENT, LEARNED SOCIETIES, AND OTHER ORGANIZATIONS

Buros, Oscar Krisen (ed.). The Fifth Mental Measurements Yearbook. Highland Park, New Jersey: The Gryphon Press, 1959.

Cornell, Ethel L. "Effects of Ability Grouping Determinable from Published Studies," The Grouping of Pupils, p. 304. Thirty-Fifth Yearbook of the National Society for the Study of Education. Bloomington, Illinois: Public School Publishing Company, 1936.

Department of Education, Province of Alberta. Curriculum Guide for Junior High School Literature. Edmonton: 1951.

Division of Curriculum, Department of Education, Province of British Columbia. Programme of Studies for Elementary Schools, Supplement, Grade VII. Victoria: Queen's Printer, 1961.

Dominion Bureau of Statistics. Ninth Census of Canada: 1951. Earnings and Employment of Wage Earners, Vol. V. Ottawa: Queen's Printer, 1953.

_____. Canada Year Book. Ottawa: Queen's Printer, 1963-64.

Wrightstone, J. Wayne. "Class Organization for Instruction," What Research Says to the Teacher, No. 13, pp. 1-33. Department of Classroom Teachers and American Educational Research Association. Washington, D.C.: National Education Association, 1957.

C. PERIODICALS

Artley, A. S. "Eclectic Approach to Reading," Elementary English, 38: 320-7, May, 1961.

- Attwell, Arthur A. and Thomas E. Linton, "Problems Inherent in Ability Grouping," Alberta Journal of Educational Research, 10:39, March, 1964.
- Baker, H. S. "Second Thoughts on Grouping," A.T.A. Magazine, 39:6-10, 34, March, 1959.
- Baldauf, Robert J. "A Comparison of the Extent of Educational Growth of Mentally Advanced Pupils in the Cedar Rapids Experiment," Journal of Educational Research, 52:181-83, January, 1959.
- Barbe, Walter B. and Tina S. Waterhouse, "An Experimental Program in Reading," Elementary English, 33:102-04, February, 1956.
- Barthelmess, Harriet, and P. Boyer, "An Evaluation of Ability Grouping," Journal of Educational Research, 26:284-94, December, 1932.
- Cochran, John R. "Grouping Students in Junior High School," Educational Leadership, 18:414-19, April, 1961.
- Dockerell, W. B. "Special Education for Gifted Children," Canadian Education and Research Digest, 2:37-44, March, 1962.
- Dolch, E. W. "Individualized Reading vs. Group Reading," Elementary English, 38:566-75, December, 1961.
- Eash, Maurice J. "Grouping: What We Have Learned," Educational Leadership, 18:430, April, 1961.
- Eckstrom, Ruth B. "Experimental Studies of Homogeneous Grouping: A Critical Review," School Review, 69:216, Summer, 1961.
- Floyd, Cecil. "Meeting Children's Reading Needs in the Middle Grades: A Preliminary Report," Elementary School Journal, 15:99-103, October, 1954.
- Guiler, W. S. and Vernon Edwards. "An Experimental Study of Methods of Instruction in Computational Arithmetic," Elementary School Journal, 43:353-60, February, 1943.
- Hall, L. "A Review of Acceleration Procedures for Regular and Special Classes," Alberta Journal of Educational Research, 4:50-57, March, 1958.
- _____. "A Review of Enrichment Procedures for Gifted Pupils," Alberta Journal of Educational Research, 4:117, June, 1958.
- Hart, R. H. "Effectiveness of an Approach to the Problem of Varying Abilities in Teaching Reading," Journal of Educational Research, 52:228-31, February, 1959.

- Holmes, Darrell, and Lois Harvey. "An Evaluation of Two Methods of Grouping," Educational Research Bulletin, 35:213-22, November, 1956.
- Jones, Daisy M. "An Experiment in Adaptation to Individual Differences," Journal of Educational Psychology, 39:257-71, May, 1948.
- Koontz, W. F. "Study of Achievement as a Function of Homogeneous Grouping," Journal of Experimental Education, 30:249-53, December, 1961.
- Livingstone, Hugh A. "Does Departmental Organization Affect Children's Adjustment?" Elementary School Journal, 61:217-20, January, 1961.
- Luchins, Abraham S. and E. H. Luchins. "Children's Attitudes Toward Homogeneous Grouping," The Journal of Genetic Psychology, 72:3-9, March, 1948.
- Mann, Maxine. "What Does Ability Grouping Do To the Self-Concept?" Childhood Education, 36:357-60, April, 1960.
- Miller, W. S. and Henry J. Otto. "Analysis of Experimental Studies in Homogeneous Grouping," Journal of Educational Research, 21:100, May, 1930.
- Morgan, Elmer F. Jr., and Gerald R. Stucker. "The Joplin Plan of Reading vs. A Traditional Method," Journal of Educational Psychology, 51:69-73, April, 1960.
- Passow, Harry A. "The Maze of Research on Ability Grouping," Educational Forum, 26:282, March, 1962.
- Russell, D. H. "Inter-Class Grouping for Reading Instruction in the Intermediate Grades," Journal of Educational Research, 39:62-70, February, 1946.
- Safford, Alton L. "Evaluation of an Individualized Reading Program," Reading Teacher, 13:266-70, April, 1960.
- Sartain, Harry W. "The Roseville Experiment with Individualized Reading," Reading Teacher, 13:277-81, April, 1960.
- Shane, H. G. "Grouping in the Elementary School," Phi Delta Kappan, 41:313-19, April, 1960.
- Wallen, Norman E. and R. O. Vowles. "The Effect of Intraclass Ability Grouping on Arithmetic Achievement in the Sixth Grade," Journal of Educational Psychology, 51:159-63, June, 1960.
- West, J. and C. Sievers. "Experiment in Cross-Grouping," Journal of Educational Research, 54:70-2, October, 1960.

- Wilhelms, F. T., and Dorothy Westby-Gibson. "Grouping Research Offers Leads," Educational Leadership, 18:412, April, 1961.
- Wilson, F. T. "Evidence About Acceleration of Gifted Youth," School and Society, 73:409-10, June, 1951.
- Witty, Paul. "Individualized Reading--A Summary and Evaluation," Elementary English, 36:401-12, 450, October, 1959.
- _____. "Nature and Extent of Educational Provisions for the Gifted Pupil," Educational Administration and Supervision, 37:65-79, February, 1951.
- Worlton, J. T. "The Why of Homogeneous Classification," Elementary School Journal, 27:265-74, 1926.
- Worth, Walter H. "Elementary Education--The Critical Years," Alberta School Trustee, 34:26, December, 1964.

D. UNPUBLISHED MATERIALS

- Arbeau, A. M. "A Survey of Pupil Grouping Practices in Grade I to XII in Alberta Schools." Unpublished Master's thesis, The University of Alberta, Edmonton, 1963.
- Melnychuk, Rudolph Steve. "Academic Achievement of Pupils in the Edmonton Continuous Progress Plan." Unpublished Master's thesis, The University of Alberta, Edmonton, 1964.
- Parker, Don H. "When Should I Individualize Instruction?" Chicago: Science Research Associates, Inc., 1963. (Mimeographed.)

B29831